



IFW16

RAW SEQUENCE LISTING

DATE: 10/07/2004

PATENT APPLICATION: US/10/005,647A

TIME: 10:08:34

Input Set : A:\sequence listing ascii 2.txt

Output Set: N:\CRF4\10072004\J005647A.raw

3 <110> APPLICANT: Jalkanen, Markku
 4 Darwish, Kamel El
 5 Lindahl, Ulf
 6 Li, Jin-Ping
 8 <120> TITLE OF INVENTION: Glucuronyl C5-Epimerase, DNA Encoding the Same and Uses
 Thereof
 10 <130> FILE REFERENCE: 1708.0280002
 12 <140> CURRENT APPLICATION NUMBER: US 10/005,647A
 13 <141> CURRENT FILING DATE: 2001-12-07
 15 <150> PRIOR APPLICATION NUMBER: US 60/304,180
 16 <151> PRIOR FILING DATE: 2000-12-08
 18 <150> PRIOR APPLICATION NUMBER: US 09/732,026
 19 <151> PRIOR FILING DATE: 2000-12-08
 21 <160> NUMBER OF SEQ ID NOS: 11
 23 <170> SOFTWARE: PatentIn version 3.2
 25 <210> SEQ ID NO: 1
 27 <211> LENGTH: 1854
 29 <212> TYPE: DNA
 31 <213> ORGANISM: Mus musculus
 35 <220> FEATURE:
 37 <221> NAME/KEY: CDS
 39 <222> LOCATION: (1)..(1854)
 43 <400> SEQUENCE: 1
 44 atg cgt tgt ttg gca gct cgg gtc aac tat aag act ttg att atc atc 48
 45 Met Arg Cys Leu Ala Ala Arg Val Asn Tyr Lys Thr Leu Ile Ile Ile
 46 1 5 10 15
 48 tgt gcg cta ttc act ttg gtc aca gta ctt ttg tgg aat aag tgt tcc 96
 49 Cys Ala Leu Phe Thr Leu Val Thr Val Leu Leu Trp Asn Lys Cys Ser
 50 20 25 30
 52 agc gac aaa gca atc cag ttt cct cgg cac ttg agt agt gga ttc aga 144
 53 Ser Asp Lys Ala Ile Gln Phe Pro Arg His Leu Ser Ser Gly Phe Arg
 54 35 40 45
 56 gtg gat gga tta gaa aaa aga tca gca gca tct gaa agt aac cac tat 192
 57 Val Asp Gly Leu Glu Lys Arg Ser Ala Ala Ser Glu Ser Asn His Tyr
 58 50 55 60
 60 gcc aac cac ata gcc aaa cag cag tca gaa gag gca ttt cct cag gaa 240
 61 Ala Asn His Ile Ala Lys Gln Gln Ser Glu Glu Ala Phe Pro Gln Glu
 62 65 70 75 80
 64 caa cag aag gca ccc cct gtt gtt ggg ggc ttc aat agc aac ggg gga 288
 65 Gln Gln Lys Ala Pro Pro Val Val Gly Gly Phe Asn Ser Asn Gly Gly
 66 85 90 95
 68 agc aag gtg tta ggg ctc aaa tat gaa gag att gac tgt ctc ata aac 336
 69 Ser Lys Val Leu Gly Leu Lys Tyr Glu Glu Ile Asp Cys Leu Ile Asn
 70 100 105 110

RAW SEQUENCE LISTING

DATE: 10/07/2004

PATENT APPLICATION: US/10/005,647A

TIME: 10:08:34

Input Set : A:\sequence listing ascii 2.txt

Output Set : N:\CRF4\10072004\J005647A.raw

72	gat	gag	cac	acc	att	aaa	ggg	aga	cga	gag	ggg	aat	gaa	gtt	ttc	ctt	384
73	Asp	Glu	His	Thr	Ile	Lys	Gly	Arg	Arg	Glu	Gly	Asn	Glu	Val	Phe	Leu	
74			115				120					125					
76	cca	ttc	act	tgg	gta	gag	aaa	tac	ttt	gat	gtt	tat	gga	aaa	gtg	gtc	432
77	Pro	Phe	Thr	Trp	Val	Glu	Lys	Tyr	Phe	Asp	Val	Tyr	Gly	Lys	Val	Val	
78		130				135						140					
80	cag	tat	gac	ggc	tat	gat	cga	ttt	gaa	ttc	tct	cat	agc	tat	tcc	aaa	480
81	Gln	Tyr	Asp	Gly	Tyr	Asp	Arg	Phe	Glu	Phe	Ser	His	Ser	Tyr	Ser	Lys	
82	145				150					155					160		
84	gtc	tat	gca	cag	aga	tca	cct	tat	cac	cct	gac	ggg	gtg	ttt	atg	tcc	528
85	Val	Tyr	Ala	Gln	Arg	Ser	Pro	Tyr	His	Pro	Asp	Gly	Val	Phe	Met	Ser	
86			165				170					175					
88	ttt	gaa	ggc	tac	aat	gtg	gaa	gtc	cga	gac	aga	gtc	aaa	tgt	ata	agt	576
89	Phe	Glu	Gly	Tyr	Asn	Val	Glu	Val	Arg	Asp	Arg	Val	Lys	Cys	Ile	Ser	
90			180				185					190					
92	gga	gtt	gaa	ggg	gtg	cca	tta	tct	acc	cag	tgg	ggg	cct	caa	ggc	tat	624
93	Gly	Val	Glu	Gly	Val	Pro	Leu	Ser	Thr	Gln	Trp	Gly	Pro	Gln	Gly	Tyr	
94		195				200						205					
96	ttc	tac	cca	atc	cag	att	gca	cag	tat	ggg	cta	agt	cat	tac	agc	aag	672
97	Phe	Tyr	Pro	Ile	Gln	Ile	Ala	Gln	Tyr	Gly	Leu	Ser	His	Tyr	Ser	Lys	
98		210				215						220					
100	aat	cta	acc	gag	aaa	ccc	cct	cac	ata	gaa	gta	tat	gaa	aca	gca	gaa	720
101	Asn	Leu	Thr	Glu	Lys	Pro	Pro	His	Ile	Glu	Val	Tyr	Glu	Thr	Ala	Glu	
102	225				230					235					240		
104	gac	agg	gac	aga	aac	atc	aga	cct	aat	gaa	tgg	act	gtg	ccc	aag	ggg	768
105	Asp	Arg	Asp	Arg	Asn	Ile	Arg	Pro	Asn	Glu	Trp	Thr	Val	Pro	Lys	Gly	
106			245				250					255					
108	tgc	ttc	atg	gcc	agt	gtg	gca	gac	aag	tct	aga	tcc	acc	aat	gtt	aaa	816
109	Cys	Phe	Met	Ala	Ser	Val	Ala	Asp	Lys	Ser	Arg	Ser	Thr	Asn	Val	Lys	
110			260				265					270					
112	cag	ttt	att	gct	cca	gaa	acc	agt	gaa	ggg	gtg	tct	ttg	cag	ctg	gga	864
113	Gln	Phe	Ile	Ala	Pro	Glu	Thr	Ser	Glu	Gly	Val	Ser	Leu	Gln	Leu	Gly	
114		275				280						285					
116	aac	aca	aaa	gac	ttc	att	att	tca	ttt	gac	ctc	aag	ctt	tta	aca	aat	912
117	Asn	Thr	Lys	Asp	Phe	Ile	Ile	Ser	Phe	Asp	Leu	Lys	Leu	Leu	Thr	Asn	
118		290				295						300					
120	ggg	agt	gtg	tct	gtg	gtt	ctg	gag	acc	aca	gaa	aag	aat	cag	ctc	ttc	960
121	Gly	Ser	Val	Ser	Val	Val	Leu	Glu	Thr	Thr	Glu	Lys	Asn	Gln	Leu	Phe	
122	305				310					315					320		
124	act	gtg	cat	tat	gtc	tca	aac	acc	cag	ctg	att	gct	ttc	aga	gac	agg	1008
125	Thr	Val	His	Tyr	Val	Ser	Asn	Thr	Gln	Leu	Ile	Ala	Phe	Arg	Asp	Arg	
126			325				330					335					
128	gac	ata	tac	tac	ggc	att	ggg	ccc	aga	act	tca	tgg	agt	aca	gtt	acc	1056
129	Asp	Ile	Tyr	Tyr	Gly	Ile	Gly	Pro	Arg	Thr	Ser	Trp	Ser	Thr	Val	Thr	
130			340				345					350					
132	aga	gac	ctg	gtc	act	gac	ctc	agg	aaa	gga	gtg	ggc	ctt	tct	aac	aca	1104
133	Arg	Asp	Leu	Val	Thr	Asp	Leu	Arg	Lys	Gly	Val	Gly	Leu	Ser	Asn	Thr	
134		355				360						365					
136	aaa	gct	gtc	aag	cca	acc	aaa	atc	atg	ccc	aaa	aag	gtg	gtt	agg	ttg	1152

RAW SEQUENCE LISTING

DATE: 10/07/2004

PATENT APPLICATION: US/10/005,647A

TIME: 10:08:34

Input Set : A:\sequence listing ascii 2.txt

Output Set: N:\CRF4\10072004\J005647A.raw

```

137 Lys Ala Val Lys Pro Thr Lys Ile Met Pro Lys Lys Val Val Arg Leu
138      370      375      380
140 att gca aaa ggg aag gga ttc ctg gac aac att acc atc tca acc aca      1200
141 Ile Ala Lys Gly Lys Gly Phe Leu Asp Asn Ile Thr Ile Ser Thr Thr
142 385      390      395      400
144 gcc cac atg gct gca ttc ttt gct gca agt gac tgg cta gtg agg aac      1248
145 Ala His Met Ala Ala Phe Phe Ala Ala Ser Asp Trp Leu Val Arg Asn
146      405      410      415
148 cag gat gag aaa ggt ggc tgg cca att atg gtg acc cgg aag tta ggg      1296
149 Gln Asp Glu Lys Gly Gly Trp Pro Ile Met Val Thr Arg Lys Leu Gly
150      420      425      430
152 gaa ggg ttt aaa tct tta gaa cca gga tgg tac tct gcc atg gca caa      1344
153 Glu Gly Phe Lys Ser Leu Glu Pro Gly Trp Tyr Ser Ala Met Ala Gln
154      435      440      445
156 ggg caa gcc atc tct acc tta gtc agg gcc tat ctt cta acg aaa gac      1392
157 Gly Gln Ala Ile Ser Thr Leu Val Arg Ala Tyr Leu Leu Thr Lys Asp
158      450      455      460
160 tat gta ttc ctc agt tca gct tta agg gca aca gcc cca tac aag ttt      1440
161 Tyr Val Phe Leu Ser Ser Ala Leu Arg Ala Thr Ala Pro Tyr Lys Phe
162 465      470      475      480
164 ccg tca gag cag cat gga gtt aaa gcc gtg ttc atg aat aaa cat gac      1488
165 Pro Ser Glu Gln His Gly Val Lys Ala Val Phe Met Asn Lys His Asp
166      485      490      495
168 tgg tat gaa gaa tat cca acc aca cct agc tct ttt gtt tta aat ggc      1536
169 Trp Tyr Glu Glu Tyr Pro Thr Thr Pro Ser Ser Phe Val Leu Asn Gly
170      500      505      510
172 ttt atg tat tct tta att ggg ctg tat gac cta aaa gaa aca gca ggg      1584
173 Phe Met Tyr Ser Leu Ile Gly Leu Tyr Asp Leu Lys Glu Thr Ala Gly
174      515      520      525
176 gag aca ctt ggg aaa gaa gca agg tcc ttg tac gag cgc ggc atg gaa      1632
177 Glu Thr Leu Gly Lys Glu Ala Arg Ser Leu Tyr Glu Arg Gly Met Glu
178      530      535      540
180 tct ctt aaa gcc atg ctg ccc ttg tat gat act ggc tcc ggg acc atc      1680
181 Ser Leu Lys Ala Met Leu Pro Leu Tyr Asp Thr Gly Ser Gly Thr Ile
182 545      550      555      560
184 tat gac ctc cgc cac ttc atg ctt ggc att gct ccc aac ctg gcc cgc      1728
185 Tyr Asp Leu Arg His Phe Met Leu Gly Ile Ala Pro Asn Leu Ala Arg
186      565      570      575
188 tgg gac tat cac acc acc cac att aac cag ctg cag ctg ctc agc acc      1776
189 Trp Asp Tyr His Thr Thr His Ile Asn Gln Leu Gln Leu Leu Ser Thr
190      580      585      590
192 atc gat gag tcc cca atc ttc aaa gaa ttt gtc aag agg tgg aaa agc      1824
193 Ile Asp Glu Ser Pro Ile Phe Lys Glu Phe Val Lys Arg Trp Lys Ser
194      595      600      605
196 tac ctt aaa ggc agt agg gca aag cac aac      1854
197 Tyr Leu Lys Gly Ser Arg Ala Lys His Asn
198      610      615
201 <210> SEQ ID NO: 2
203 <211> LENGTH: 618

```

RAW SEQUENCE LISTING

DATE: 10/07/2004

PATENT APPLICATION: US/10/005,647A

TIME: 10:08:34

Input Set : A:\sequence listing ascii 2.txt

Output Set: N:\CRF4\10072004\J005647A.raw

```

205 <212> TYPE: PRT
207 <213> ORGANISM: Mus musculus
211 <400> SEQUENCE: 2
213 Met Arg Cys Leu Ala Ala Arg Val Asn Tyr Lys Thr Leu Ile Ile Ile
214 1 5 10 15
217 Cys Ala Leu Phe Thr Leu Val Thr Val Leu Leu Trp Asn Lys Cys Ser
218 20 25 30
221 Ser Asp Lys Ala Ile Gln Phe Pro Arg His Leu Ser Ser Gly Phe Arg
222 35 40 45
225 Val Asp Gly Leu Glu Lys Arg Ser Ala Ala Ser Glu Ser Asn His Tyr
226 50 55 60
229 Ala Asn His Ile Ala Lys Gln Gln Ser Glu Glu Ala Phe Pro Gln Glu
230 65 70 75 80
233 Gln Gln Lys Ala Pro Pro Val Val Gly Gly Phe Asn Ser Asn Gly Gly
234 85 90 95
237 Ser Lys Val Leu Gly Leu Lys Tyr Glu Glu Ile Asp Cys Leu Ile Asn
238 100 105 110
241 Asp Glu His Thr Ile Lys Gly Arg Arg Glu Gly Asn Glu Val Phe Leu
242 115 120 125
245 Pro Phe Thr Trp Val Glu Lys Tyr Phe Asp Val Tyr Gly Lys Val Val
246 130 135 140
249 Gln Tyr Asp Gly Tyr Asp Arg Phe Glu Phe Ser His Ser Tyr Ser Lys
250 145 150 155 160
253 Val Tyr Ala Gln Arg Ser Pro Tyr His Pro Asp Gly Val Phe Met Ser
254 165 170 175
257 Phe Glu Gly Tyr Asn Val Glu Val Arg Asp Arg Val Lys Cys Ile Ser
258 180 185 190
261 Gly Val Glu Gly Val Pro Leu Ser Thr Gln Trp Gly Pro Gln Gly Tyr
262 195 200 205
265 Phe Tyr Pro Ile Gln Ile Ala Gln Tyr Gly Leu Ser His Tyr Ser Lys
266 210 215 220
269 Asn Leu Thr Glu Lys Pro Pro His Ile Glu Val Tyr Glu Thr Ala Glu
270 225 230 235 240
273 Asp Arg Asp Arg Asn Ile Arg Pro Asn Glu Trp Thr Val Pro Lys Gly
274 245 250 255
277 Cys Phe Met Ala Ser Val Ala Asp Lys Ser Arg Ser Thr Asn Val Lys
278 260 265 270
281 Gln Phe Ile Ala Pro Glu Thr Ser Glu Gly Val Ser Leu Gln Leu Gly
282 275 280 285
285 Asn Thr Lys Asp Phe Ile Ile Ser Phe Asp Leu Lys Leu Leu Thr Asn
286 290 295 300
289 Gly Ser Val Ser Val Val Leu Glu Thr Thr Glu Lys Asn Gln Leu Phe
290 305 310 315 320
293 Thr Val His Tyr Val Ser Asn Thr Gln Leu Ile Ala Phe Arg Asp Arg
294 325 330 335
297 Asp Ile Tyr Tyr Gly Ile Gly Pro Arg Thr Ser Trp Ser Thr Val Thr
298 340 345 350
301 Arg Asp Leu Val Thr Asp Leu Arg Lys Gly Val Gly Leu Ser Asn Thr
302 355 360 365

```

RAW SEQUENCE LISTING

DATE: 10/07/2004

PATENT APPLICATION: US/10/005,647A

TIME: 10:08:34

Input Set : A:\sequence listing ascii 2.txt

Output Set: N:\CRF4\10072004\J005647A.raw

```

305 Lys Ala Val Lys Pro Thr Lys Ile Met Pro Lys Lys Val Val Arg Leu
306      370      375      380
309 Ile Ala Lys Gly Lys Gly Phe Leu Asp Asn Ile Thr Ile Ser Thr Thr
310 385      390      395      400
313 Ala His Met Ala Ala Phe Phe Ala Ala Ser Asp Trp Leu Val Arg Asn
314      405      410      415
317 Gln Asp Glu Lys Gly Gly Trp Pro Ile Met Val Thr Arg Lys Leu Gly
318      420      425      430
321 Glu Gly Phe Lys Ser Leu Glu Pro Gly Trp Tyr Ser Ala Met Ala Gln
322      435      440      445
325 Gly Gln Ala Ile Ser Thr Leu Val Arg Ala Tyr Leu Leu Thr Lys Asp
326      450      455      460
329 Tyr Val Phe Leu Ser Ser Ala Leu Arg Ala Thr Ala Pro Tyr Lys Phe
330 465      470      475      480
333 Pro Ser Glu Gln His Gly Val Lys Ala Val Phe Met Asn Lys His Asp
334      485      490      495
337 Trp Tyr Glu Glu Tyr Pro Thr Thr Pro Ser Ser Phe Val Leu Asn Gly
338      500      505      510
341 Phe Met Tyr Ser Leu Ile Gly Leu Tyr Asp Leu Lys Glu Thr Ala Gly
342      515      520      525
345 Glu Thr Leu Gly Lys Glu Ala Arg Ser Leu Tyr Glu Arg Gly Met Glu
346      530      535      540
349 Ser Leu Lys Ala Met Leu Pro Leu Tyr Asp Thr Gly Ser Gly Thr Ile
350 545      550      555      560
353 Tyr Asp Leu Arg His Phe Met Leu Gly Ile Ala Pro Asn Leu Ala Arg
354      565      570      575
357 Trp Asp Tyr His Thr Thr His Ile Asn Gln Leu Gln Leu Leu Ser Thr
358      580      585      590
361 Ile Asp Glu Ser Pro Ile Phe Lys Glu Phe Val Lys Arg Trp Lys Ser
362      595      600      605
365 Tyr Leu Lys Gly Ser Arg Ala Lys His Asn
366      610      615

```

369 <210> SEQ ID NO: 3

370 <211> LENGTH: 2239

371 <212> TYPE: DNA

372 <213> ORGANISM: Artificial sequence

374 <220> FEATURE:

375 <223> OTHER INFORMATION: Fusion protein having the sequence of the bovine C5-epimerase and

376 the N-terminus of the mouse C5-epimerase

378 <400> SEQUENCE: 3

```

379 tgggtgtcct ggaactcact ctgtagacca ggctggccat gaactcacag agatctacct      60
381 cctgagtgtc gggattaaag gtttgtgcca ccacctccca actctaaggt gtttctttaa      120
383 gttaggggca tagtaaacat tggtgagata ctagaggaac actgaatgaa aatttgga      180
385 tctctgcttt aggtttgtgc tgagcagttt gcctcttata ttcacctatg ctgaaaagtt      240
387 tgagttcata attttgaaca tgcatatgat aaaatattct ggccgcacat tgaataaata      300
389 tatttttaaat gaacttacct ttaaaatgtc agtaacaact ctgcatgggt ttcttcttac      360
391 ctccataggt atggtctgaa tatgcgttgt ttggcagctc gggccaacta taagactttg      420
393 attatcatct gtgcgtatt cactttggtc acagtacttt tgtggaataa gtgttccagc      480
395 gacaaagcaa tccagtttcc tcggcacttg agtagtggaat tcagagtgga tggattagaa      540

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/005,647A

DATE: 10/07/2004

TIME: 10:08:35

Input Set : A:\sequence listing ascii 2.txt

Output Set: N:\CRF4\10072004\J005647A.raw